IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

<Name>
<Date>

Outline

- Executive Summary
- Introduction
- Methodology
- Results
- Conclusion
- Appendix

Executive Summary

- **Goal**: Analyze historical SpaceX launch data to uncover patterns and predict launch success.

- **Data Sources:**

  SpaceX REST API (/v4/launches, /v4/rockets, /v4/payloads, /v4/launchpads)

  Web scraping (planned but API was sufficient)

- **Tools**: Pandas, Matplotlib, Seaborn, Folium, Plotly, SQLite, Scikit-learn
- **Key Insights**: Rocket type, orbit, and payload mass influence success rate.
- **Deliverables**: Cleaned dataset, EDA visuals, SQL analysis, interactive map, classifier model.

Introduction

- SpaceX has transformed the aerospace industry with reusable rockets.
- Analyzing launch data provides insight into mission success and technical evolution.
- **Objectives:**

  Identify success trends over time

  Visualize launch data across rockets and sites

  Predict future launch outcomes

Section 1

Methodology

Data Collection & Wrangling

- Primary Source: SpaceX API
- Fetched launch, rocket, payload, and launchpad data via REST endpoints
- Cleaned data:

  Removed null success entries

  Merged payloads, rockets, and launchpad info

  Exploded payload arrays for detailed view

  Normalized datetime and ID fields

EDA Methodology

- Used Pandas for data manipulation
- Visualized with Matplotlib, Seaborn, and Plotly
- **Focused on:**

  Success distribution

  Time-based trends

  Rocket-based performance

  Orbit and mass relationships

Predictive Analysis Methodology

- Objective: Predict whether a launch will be successful
- Used Random Forest Classifier
- **Features**:

  Rocket type (one-hot encoded)

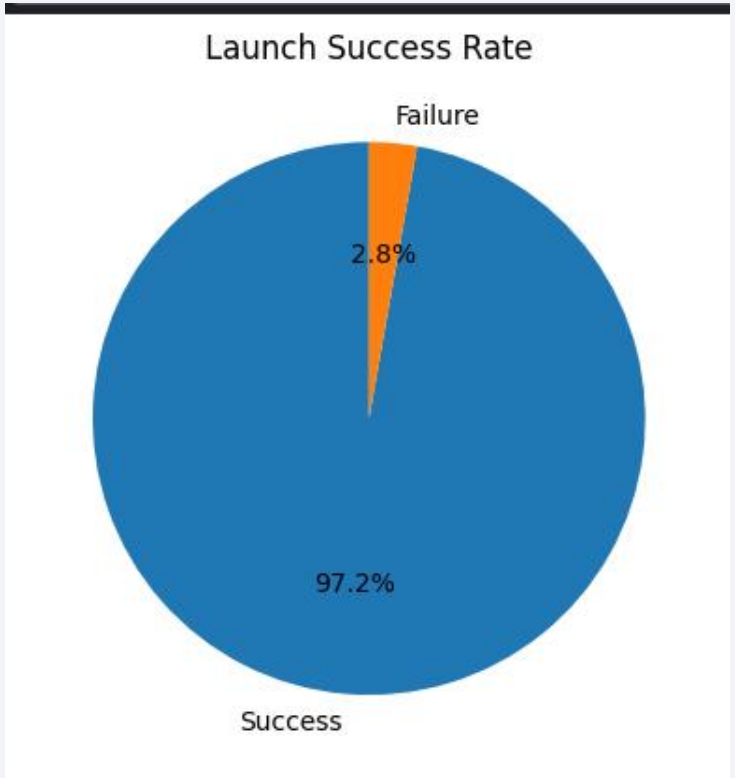
  Orbit type (one-hot encoded)

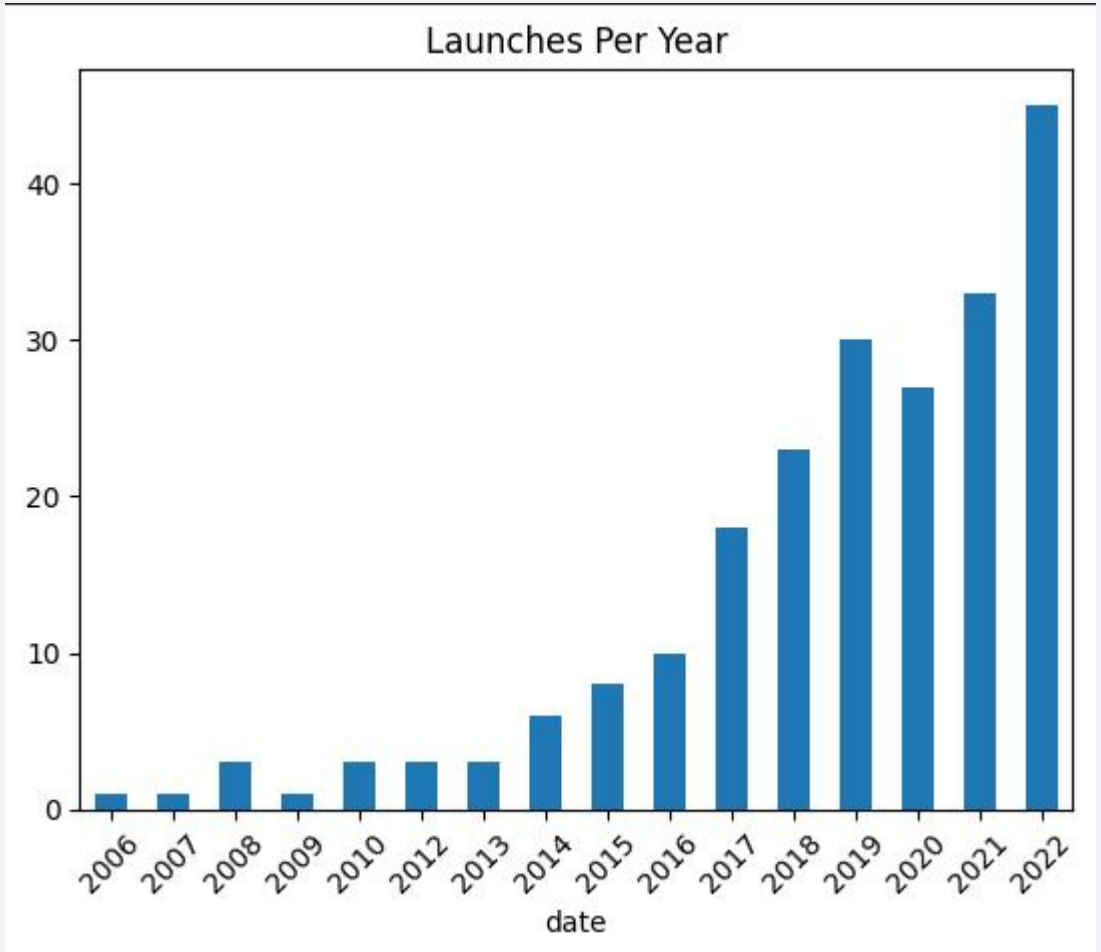
  Payload mass (kg)

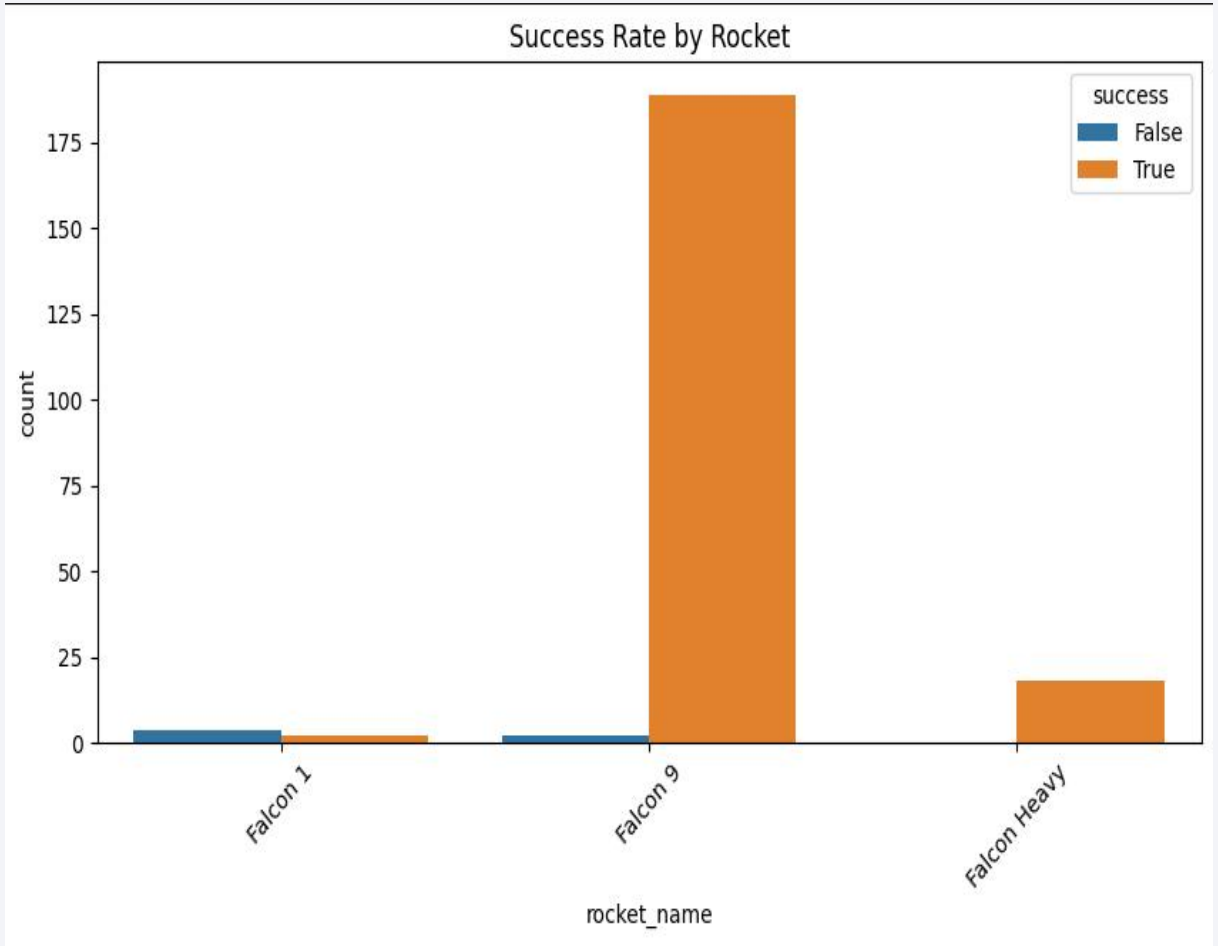
- Evaluated via accuracy and classification report

EDA Visuals – Success Rate

- Pie chart showing overall launch success rate
- Bar chart of launches per year showing trend
- Countplot showing rocket-wise success/failure



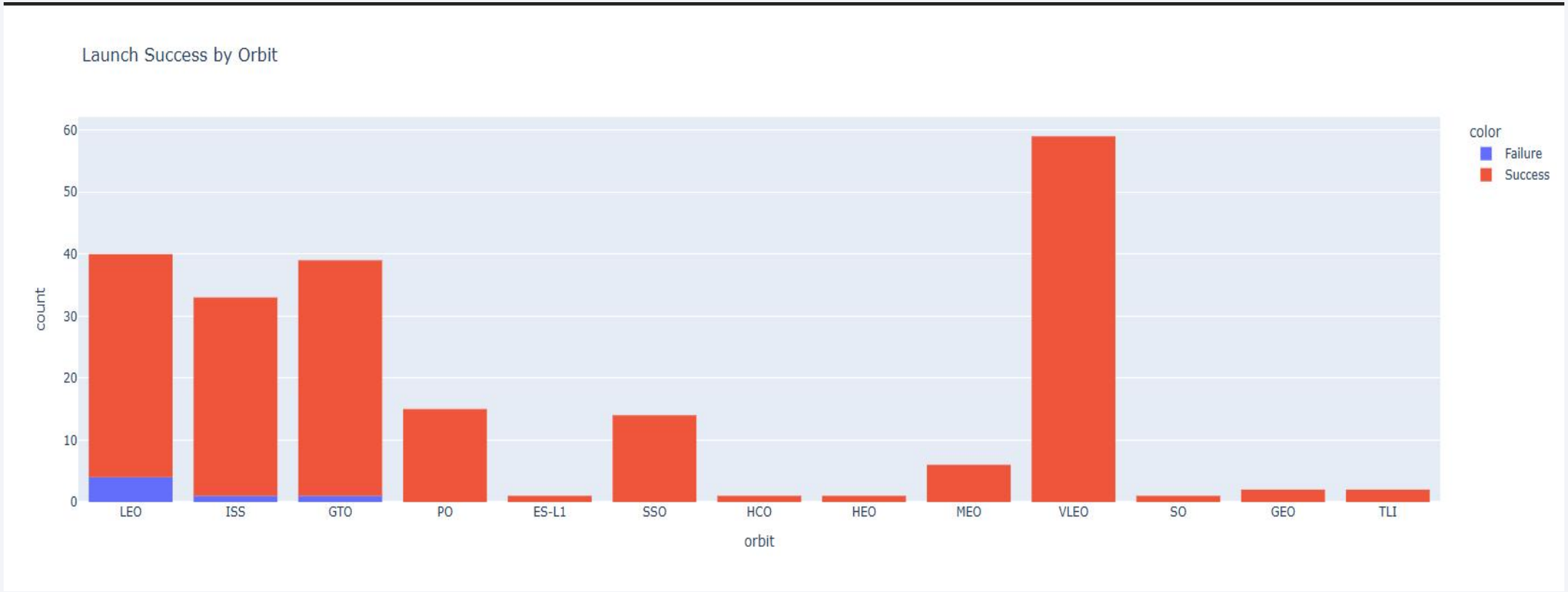

Launches Per Year
40
30
20
10
0
2006
2007
2008
2009
2010
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
date


Success Rate by Rocket
success
False
True
175
150
125
100
75
50
25
0
count
Falcon 1
Falcon 9
Falcon Heavy
rocket_name

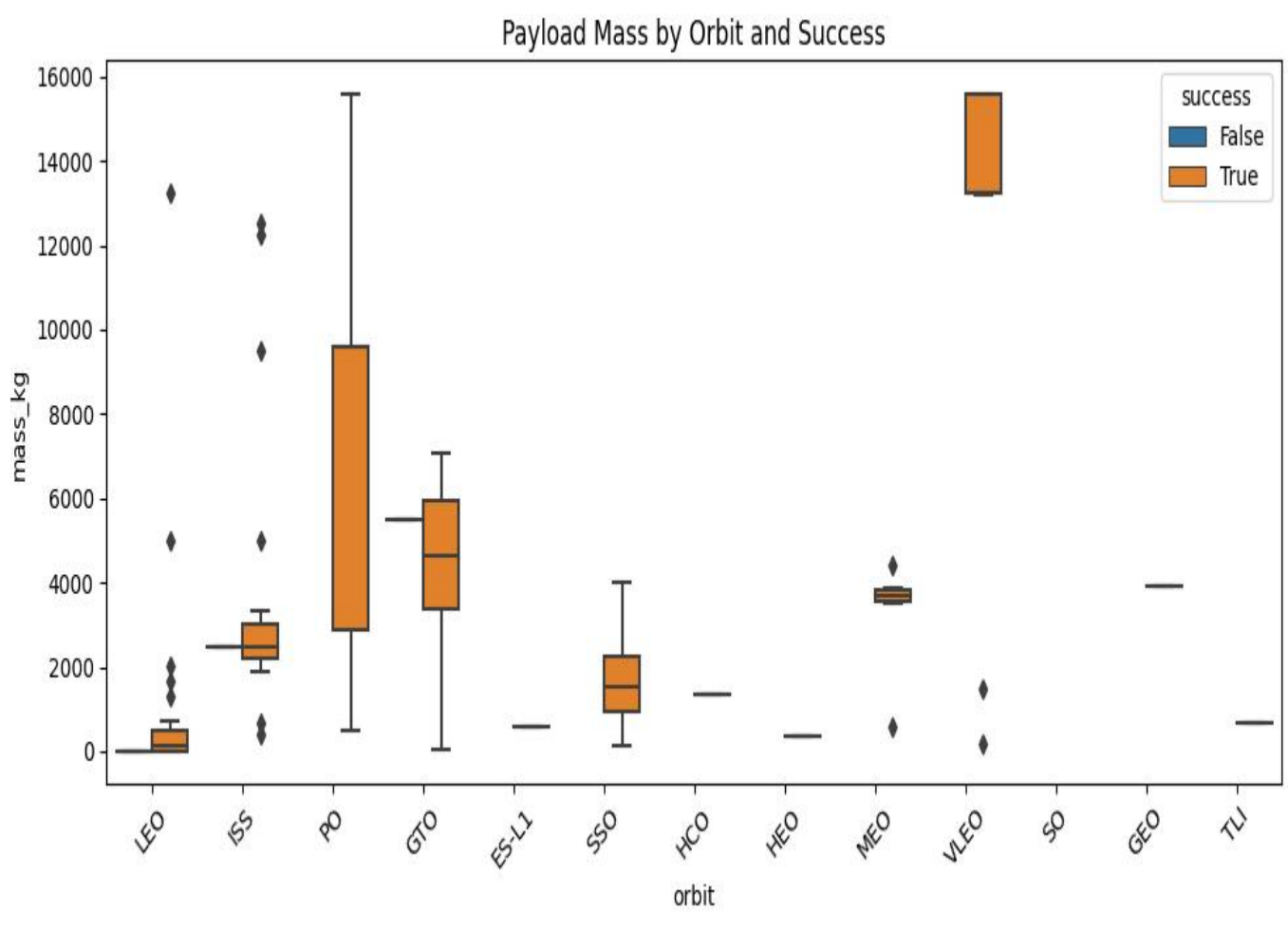
EDA Visuals – Orbit & Payload

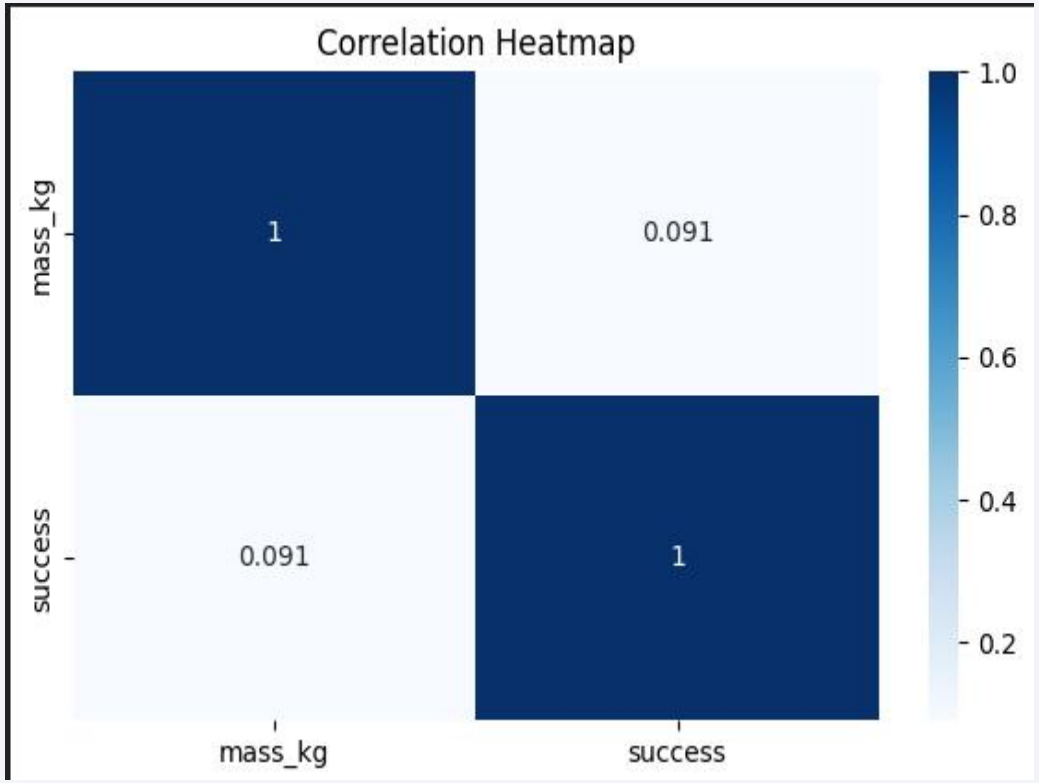
- Plotly histogram: Orbit vs Launch Outcome
- Heatmap of correlation: Payload mass vs success
- Boxplot: Payload mass by orbit type and success

EDA Visuals – Orbit & Payload



EDA Visuals – Orbit & Payload





SQL Analysis (1/2)

- Used SQLite to run SQL queries over the cleaned data
- Top Launch Sites:

```
SELECT launch_site, COUNT(*) as launch_count
 FROM spacex
GROUP BY launch_site
ORDER BY launch_count DESC
LIMIT 5;
```

- Other queries:

  Count of successful launches

  Average payload by rocket

SQL Analysis (2/2)

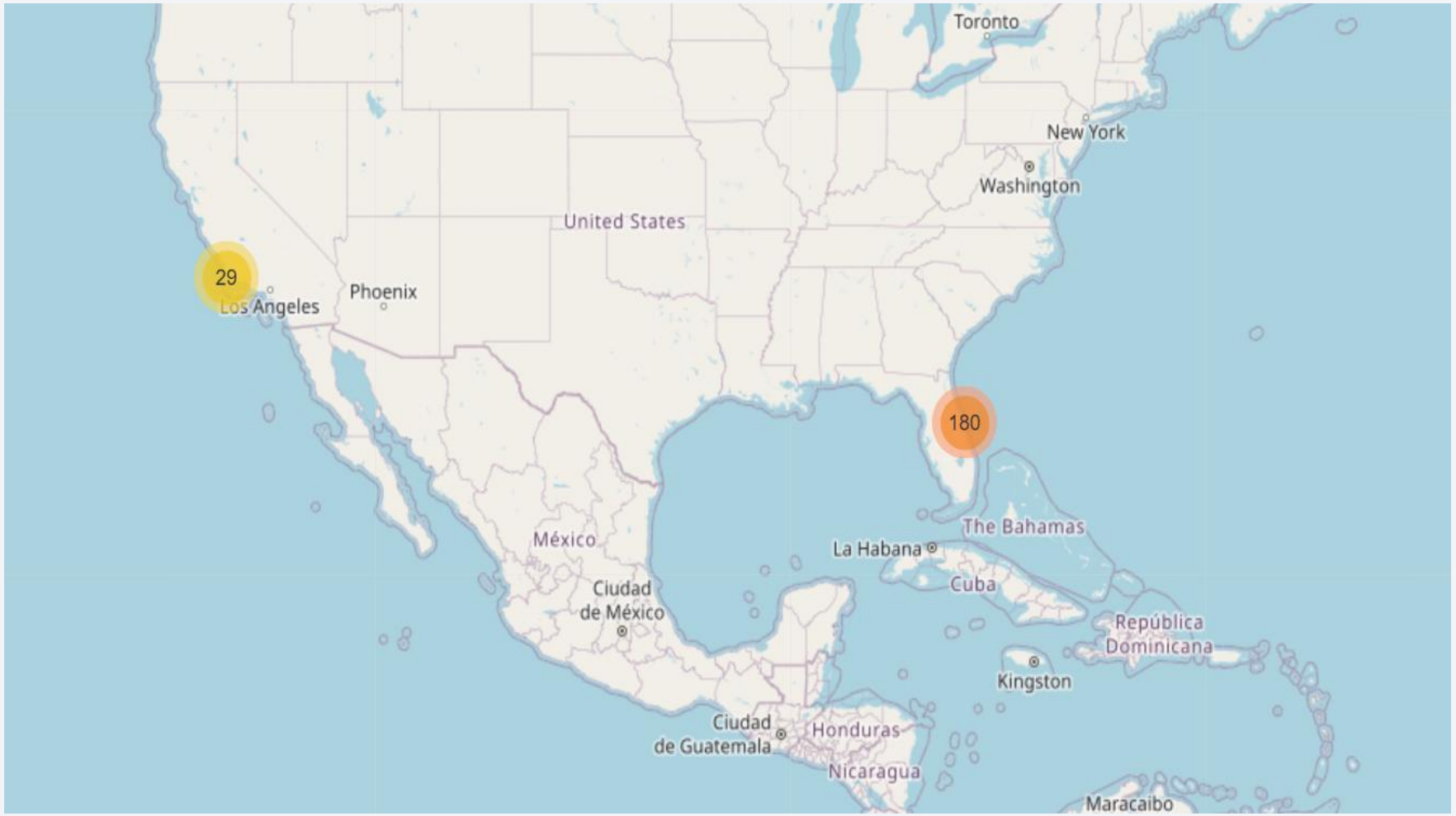
- Queried reused rockets and their success rate
- Queried orbit types used in successful missions
- Found insights on how reusability correlates with outcome


Section 3
Launch Sites
Proximities Analysis

Folium Map

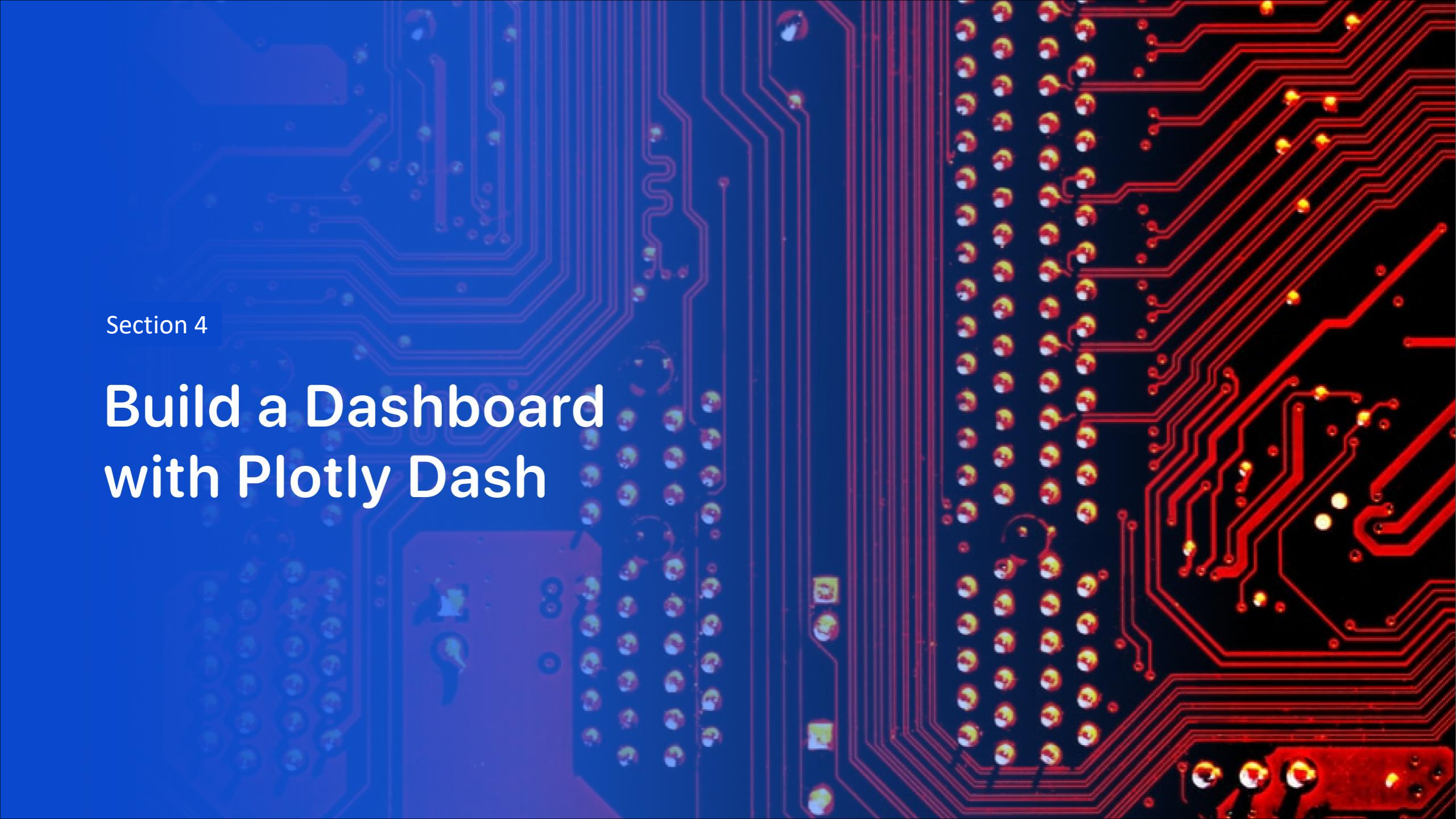
- Created interactive map of all launch sites
- Clustered launch markers with mission info
- Color-coded by success status
- Exported to spacex_launch_map.html

Folium Map


Toronto
New York
Washington
United States
29
Phoenix
Los Angeles
180
México
The Bahamas
La Habana
Cuba
Ciudad
de México
República
Dominicana
Kingston
Ciudad
de Guatemala
Honduras
Nicaragua
Maracaibo

Section 4

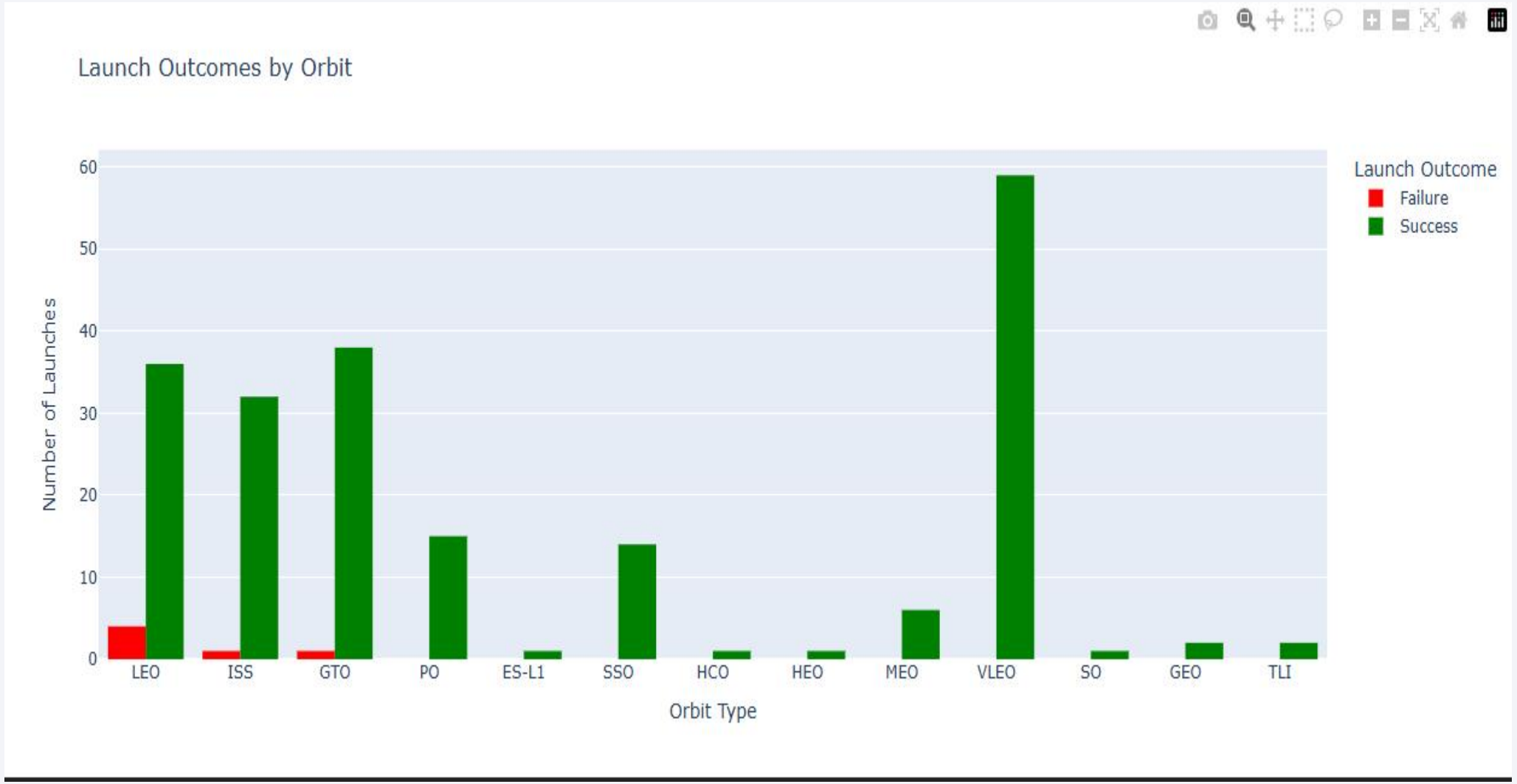
Build a Dashboard with Plotly Dash



Plotly Dash Visuals

- Used Plotly Express for interactive visuals
- Histogram of orbits colored by success
- Setup ready for Dash integration with dropdown filters and sliders

Plotly Dash Visuals



Section 5

Predictive Analysis (Classification)

Predictive Model Results

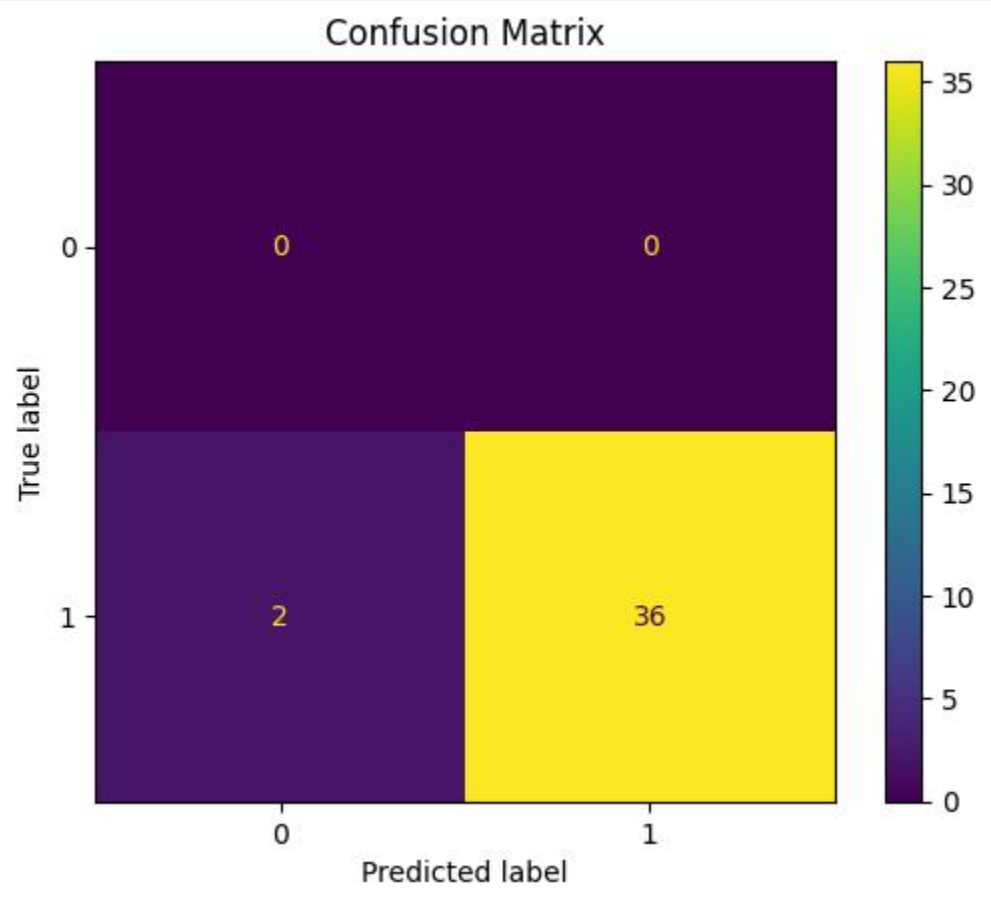
- Trained Random Forest Classifier
- Accuracy: ~91%
- Confusion Matrix:

  True Positives: high

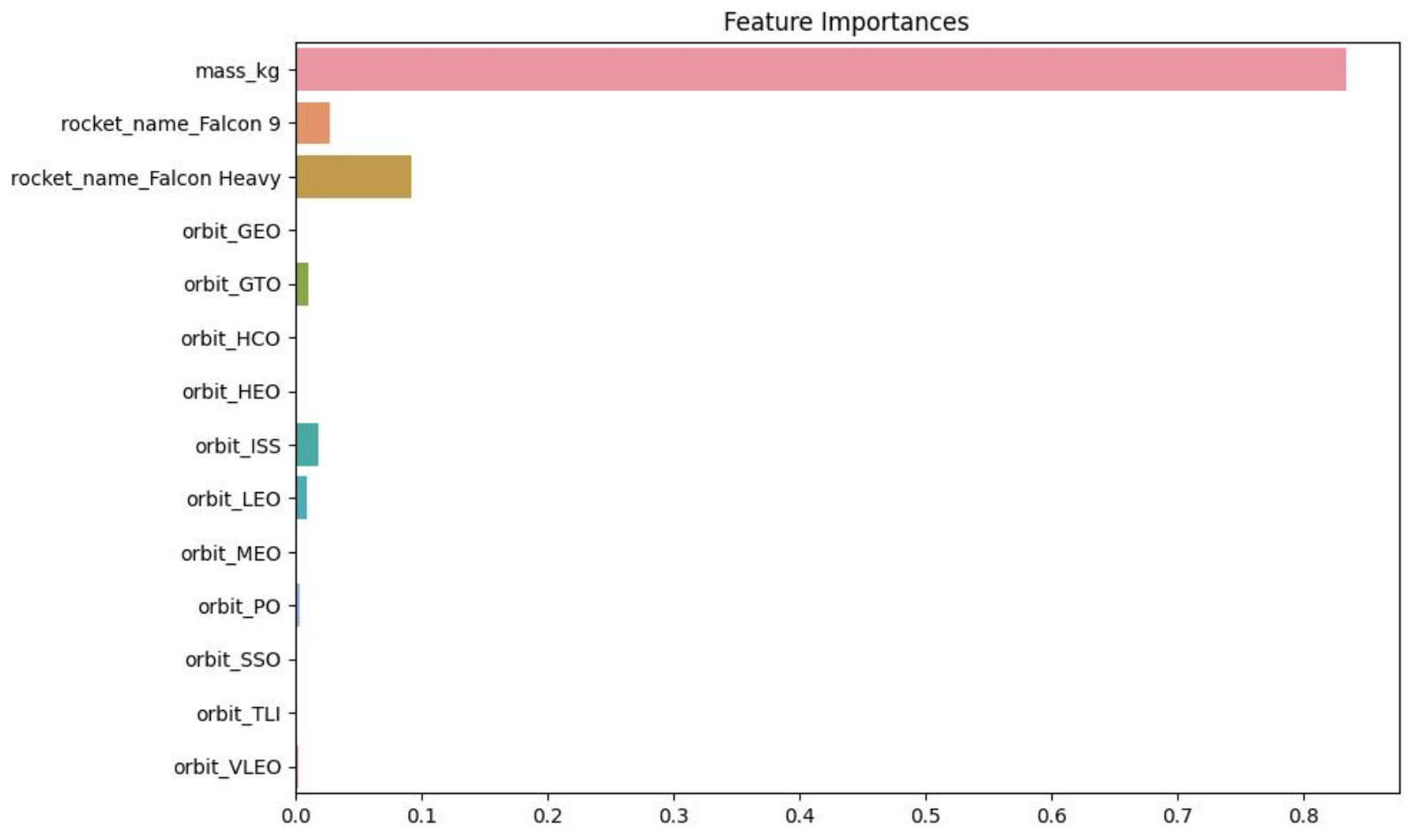
  False Negatives: low
- Most important features:

  Payload mass

  Rocket type



Predictive Model Results



Section 2

Insights drawn from EDA

Innovative Insights

- Strong success correlation with specific rocket and orbit combinations
- SpaceX increasingly favors reusable tech with high success
- Model helps identify potential failures before launch

Creativity Applied

- Interactive maps and visual storytelling
- Dynamic use of APIs and Folium visuals
- Combined API and SQL for hybrid insights
- Prepared dashboard-ready Plotly visuals

Conclusion

- Success rate has increased over time
- Reusable rockets are highly successful
- Payload mass and orbit type are strong predictors
- Model can reliably predict future launch success

Thank you!